



**QUEEN'S  
UNIVERSITY  
BELFAST**

## Optimal pilot length for uplink massive MIMO systems with low-resolutions ADCs

Fan, L., Qiao, D., Jin, S., Wen, C-K., & Matthaiou, M. (2016). Optimal pilot length for uplink massive MIMO systems with low-resolutions ADCs. In *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)* Institute of Electrical and Electronics Engineers Inc.. <https://doi.org/10.1109/SAM.2016.7569691>

### Published in:

IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)

### Document Version:

Peer reviewed version

### Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

### Publisher rights

© 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

### General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact [openaccess@qub.ac.uk](mailto:openaccess@qub.ac.uk).

# Optimal Pilot Length for Uplink Massive MIMO Systems with Low-Resolution ADC

Li Fan, Dan Qiao, Shi Jin

National Mobile Communication Research Laboratory Institute of Communications Engineering Queen's University Belfast  
Southeast University  
Nanjing, China

Chao-Kai Wen

National Sun Yat-sen University  
Kaohsiung, Taiwan

Michail Matthaiou

Northern Ireland, U.K.

**Abstract**—In this paper, we consider the problem of pilot based data transmission in massive multiple-input multiple-output (MIMO) systems where each receiver antenna is equipped with a low-precision analog-to-digital converter (ADC). Under the quantized MIMO setting, we derive a tight approximation of the uplink spectral efficiency considering a maximal ratio combining receiver. Capitalizing on the derived expression, the optimal pilot length which maximizes the sum spectral efficiency is put forward. Our results show that massive MIMO with low-resolution ADCs require more training time compared to the case with high-resolution quantizers. In particular, for the receivers with one-bit quantizers, the effective way to maximize the sum spectral efficiency is to assign more pilot symbols instead of increasing the transmit power.

**Keywords**—Massive MIMO; pilot length; quantization; spectral efficiency

## I. INTRODUCTION

MASSIVE multiple-input multiple-output (MIMO) is being investigated as a key technique for the fifth generation (5G) cellular networks since it can achieve remarkable energy and spectral efficiency (SE) [1, 2]. However, a substantial increase in hardware cost and power consumption especially from the ADC front-ends, hinder the realization of massive MIMO systems. This fact motivates the use of quantized MIMO systems where each receiver antenna uses a very low-resolution (e.g., 1–3 bits) analog-to-digital converter (ADC) [3–5]. Results in [3, 5] demonstrated that the rate degradation due to the low-resolution ADCs can be compensated by increasing the number of receiving antennas. By evaluating the effect of ADC resolution and bandwidth on the achievable rate under a receiver power constraint, [4] also advocates the quantized MIMO architecture.

However, these papers assume perfect channel state information at the receiver, which could result in an over-optimistic estimate for the quantized MIMO system. The use of coarse quantization reduces greatly the number of effective samples and hinders the channel estimate. Channel estimation is typically performed using uplink pilot training. A natural question is the following: *how many resources in terms of*

*time and power should be allocated for training in quantized massive MIMO systems?* Although partial answers were given in [6], the related literature is still very scarce.

To fill this literature gap, we investigate the uplink SE of quantized massive MIMO systems, where channel estimation is performed first followed by the maximal ratio combining (MRC). A tight approximation of the uplink SE is obtained. Capitalizing on the derived approximation, the optimal pilot length which maximizes the sum SE is derived. Our results show that massive MIMO with low-resolution ADCs require more training time.

**Notations:** Throughout the paper, boldfaced letters are used to represent vectors and matrices. The operator  $(\cdot)^H$  and  $(\cdot)^*$  stand for the Hermitian and conjugate of a matrix, respectively. Moreover,  $\mathbf{I}_N$  denotes an  $N \times N$  identity matrix, and  $\text{diag}(\mathbf{A})$  denotes the diagonal matrix with the same diagonal elements as matrix  $\mathbf{A}$ . Finally,  $\mathbb{E}\{\cdot\}$  is the expectation operator and  $\|\cdot\|$  is the Euclidean norm.

## II. SYSTEM MODEL

We consider the uplink of a multi-user MIMO system where a base station (BS) is equipped with an array of  $M$  antennas and serves  $N$  single-antenna user terminals in the same time-frequency resource. The received  $M$ -dimensional vector  $\mathbf{y}$  at the BS can be expressed as

$$\mathbf{y} = \sqrt{p} \mathbf{G} \mathbf{x} + \mathbf{n}, \quad (1)$$

where  $\mathbf{G} = [g_{mn}] \in \mathbb{C}^{M \times N}$  represents the channel matrix between the BS and terminals,  $\mathbf{x}$  denotes the  $N$ -dimensional zero-mean vector of Gaussian symbols transmitted by all users such that  $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \mathbf{I}_N$ ,  $p$  is the average transmitted power of each user, and  $\mathbf{n}$  is the additive white Gaussian noise (AWGN) vector with zero mean and unit element-wise variance written  $\mathbf{n} \sim \mathcal{CN}(0, \mathbf{I}_M)$ . The channel coefficient between the  $n$ th user terminal and the  $m$ th antenna at the BS can be modeled as  $g_{mn} = h_{mn} \sqrt{\beta_n}$ , where  $h_{mn}$  represents the small-scale fading coefficient, and  $\beta_n$  denotes the large-scale fading coefficient. Following [1], we assume that  $h_{mn} \sim \mathcal{CN}(0, 1)$ , and  $\beta_n$  is constant across the antenna array. In matrix form, we write

$$\mathbf{G} = \mathbf{H} \mathbf{D}^{1/2}, \quad (2)$$

where  $\mathbf{H} = [h_{mn}] \in \mathbb{C}^{M \times N}$ , and  $\mathbf{D}$  is a  $N \times N$  diagonal matrix with diagonal entries  $\{\beta_n\}$ .

This work was supported by the National Natural Science Foundation of China under Grant 61531011 and the International Science and Technology Cooperation Program of China under Grant 2014DFT10300. The work of C.-K. Wen was supported in part by the Ministry of Science and Technology of Taiwan under Grants MOST 103-2221-E-110-029-MY3 and the ITRI in Hsinchu, Taiwan.

Assuming that the gain of automatic gain control is set appropriately, we adopt the additive quantization noise model (AQNM) and formulate the quantizer outputs as [7]

$$\tilde{\mathbf{y}} = \alpha \mathbf{y} + \tilde{\mathbf{n}} = \alpha \sqrt{p} \mathbf{G} \mathbf{x} + \alpha \mathbf{n} + \tilde{\mathbf{n}}, \quad (3)$$

where  $\alpha = 1 - \rho \in [0, 1]$  with  $\rho$  being the inverse of the signal-to-quantization-noise ratio, and  $\tilde{\mathbf{n}}$  is the additive Gaussian quantization noise vector that is statistically uncorrelated with  $\mathbf{y}$  and its covariance matrix is given by

$$\mathbf{R}_{\tilde{\mathbf{n}}} = \alpha(1 - \alpha) \text{diag}(\mathbf{R}_{\mathbf{y}}) \quad (4)$$

with  $\mathbf{R}_{\mathbf{y}} = \mathbb{E}\{\mathbf{y}\mathbf{y}^*\}$ . Note that the values of  $\rho$  (or  $\alpha$ ) depend on the number of quantization bits  $b$ , which can be found in [5]. The uplink transmission comprises two phases: training and data transmission phases.

#### A. Training Phase

In the training phase, the users transmit a sequence of known pilot signals of symbols during the channel coherence time  $T$ . Let  $\psi_n = [\psi_{ni}] \in \mathbb{C}^{\tau \times 1}$  be the pilot vector for the  $n$ th user, with  $\tau$  being the number of training symbols, and  $\mathbf{g}_n$  be the  $n$ th column of the channel matrix  $\mathbf{G}$ . The received signal at the  $i$ th symbol can be written as

$$\mathbf{y}_i^t = \sqrt{\tau p_t} \sum_{n=1}^N \psi_{ni} \mathbf{g}_n + \mathbf{n}_i^t, \quad (5)$$

where  $p_t$  denotes the transmit power associated to each one of the pilot symbol,  $\psi_{ni}$  is the  $i$ th element of  $\psi_n$ , and  $\mathbf{n}_i^t$  is the AWGN vector during pilot transmission. Then, following (3), the quantizer outputs can be written as

$$\tilde{\mathbf{y}}_i^t = \alpha \mathbf{y}_i^t + \tilde{\mathbf{n}}_i^t = \alpha \sqrt{\tau p_t} \sum_{n=1}^N \psi_{ni} \mathbf{g}_n + \alpha \mathbf{n}_i^t + \tilde{\mathbf{n}}_i^t, \quad (6)$$

where  $\tilde{\mathbf{n}}_i^t$  is the additive Gaussian quantization noise vector whose covariance matrix is

$$\mathbf{R}_{\tilde{\mathbf{n}}_i^t} = \alpha(1 - \alpha) \left( 1 + \tau p_t \sum_{n=1}^N |\psi_{ni}|^2 \beta_n \right) \mathbf{I}_M. \quad (7)$$

From (7), we find that the covariance of  $\tilde{\mathbf{n}}_i^t$  is dependant on the pilot symbols  $\psi_{ni}$ . Here we assume that the pilot sequence  $\psi_n$  is the  $n$ th row of a discrete Fourier transform matrix, we have  $\mathbf{R}_{\tilde{\mathbf{n}}_i^t} = \tilde{\sigma}_t^2 \mathbf{I}_M$  with

$$\tilde{\sigma}_t^2 = \alpha(1 - \alpha) \left( 1 + p_t \sum_{n=1}^N \beta_n \right). \quad (8)$$

Let  $\tilde{\mathbf{Y}}^t = [\tilde{\mathbf{y}}_1^t, \dots, \tilde{\mathbf{y}}_\tau^t] \in \mathbb{C}^{M \times \tau}$ ,  $\mathbf{N}^t = [\mathbf{n}_1^t, \dots, \mathbf{n}_\tau^t] \in \mathbb{C}^{M \times \tau}$ ,  $\tilde{\mathbf{N}}^t = [\tilde{\mathbf{n}}_1^t, \dots, \tilde{\mathbf{n}}_\tau^t] \in \mathbb{C}^{M \times \tau}$ , and  $\Psi = [\psi_1^T, \dots, \psi_N^T]^T \in \mathbb{C}^{N \times \tau}$ . We can write the quantizer outputs (6) in matrix form as

$$\tilde{\mathbf{Y}}^t = \alpha \sqrt{p_t} \mathbf{G} \Psi + \alpha \mathbf{N}^t + \tilde{\mathbf{N}}^t. \quad (9)$$

The minimum mean-square-error (MMSE) estimate of  $\mathbf{G}$  is [8]

$$\hat{\mathbf{G}} = \frac{1}{\alpha \sqrt{\tau p_t}} \tilde{\mathbf{Y}}^t \Psi^H \tilde{\mathbf{D}}, \quad (10)$$

where  $\tilde{\mathbf{D}} = \left[ \mathbf{I}_N + \left( \frac{1}{\tau p_t} + \frac{\tilde{\sigma}_t^2}{\alpha^2 \tau p_t} \right) \mathbf{D}^{-1} \right]^{-1}$ . Let  $\mathbf{E} = \mathbf{G} - \hat{\mathbf{G}} = [e_{mi}]$  denote the channel estimation error matrix. After some basic calculations, we obtain

$$\sigma_{\hat{g}_i}^2 = \mathbb{E} \left\{ |e_{mi} - \mathbb{E}\{e_{mi}\}|^2 \right\} = \frac{(\alpha^2 + \tilde{\sigma}_t^2) \beta_i}{\tau \alpha^2 p_t \beta_i + \alpha^2 + \tilde{\sigma}_t^2}. \quad (11)$$

First, we note that the variance of the channel estimation error  $\sigma_{\hat{g}_i}^2$  is independent on the antenna index  $m$ . Second, we can show that  $\partial \sigma_{\hat{g}_i}^2 / \partial \alpha < 0$ . This fact indicates that  $\sigma_{\hat{g}_i}^2$  is monotonically decreasing with  $\alpha \in [0, 1]$  and reaches the minimum at  $\alpha = 1$ . This result is expected because the accuracy of channel estimate can be enhanced by increasing the number of quantization bits.

#### B. Data Transmission Phase

In the data transmission phase, all terminals send their data to the BS using the remaining  $T - \tau$  symbols. Let  $\sqrt{p_d} \mathbf{x}^d \in \mathbb{C}^{N \times 1}$  be the vector of uplink symbols transmitted from all the terminals, where  $\mathbb{E} \{ \mathbf{x}^d (\mathbf{x}^d)^H \} = \mathbf{I}_N$  and  $p_d$  is the average transmitted power during data transmission. The received vector at the BS is given by

$$\mathbf{y}^d = \sqrt{p_d} \mathbf{G} \mathbf{x}^d + \mathbf{n}^d, \quad (12)$$

where  $\mathbf{n}^d \in \mathbb{C}^{M \times 1}$  is the AWGN vector. Following (3), the quantizer output is

$$\tilde{\mathbf{y}}^d = \alpha \sqrt{p_d} \mathbf{G} \mathbf{x}^d + \alpha \mathbf{n}^d + \tilde{\mathbf{n}}^d \quad (13)$$

with  $\tilde{\mathbf{n}}^d$  being the additive Gaussian quantization noise vector. The BS uses the MRC together with the estimated channel to detect the transmitted signals as

$$\hat{\mathbf{x}}^d = \hat{\mathbf{G}}^H \tilde{\mathbf{y}}^d = \alpha \sqrt{p_d} \hat{\mathbf{G}}^H (\hat{\mathbf{G}} + \mathbf{E}) \mathbf{x}^d + \alpha \hat{\mathbf{G}}^H \mathbf{n}^d + \hat{\mathbf{G}}^H \tilde{\mathbf{n}}^d, \quad (14)$$

where the second equality follows from the fact that  $\mathbf{E} = \mathbf{G} - \hat{\mathbf{G}}$ . According to (14), the detected signal for the  $n$ th terminal is given by

$$\begin{aligned} \hat{x}_n^d &= \alpha \sqrt{p_d} \hat{\mathbf{g}}_n^H \hat{\mathbf{g}}_n x_n^d + \alpha \sqrt{p_d} \sum_{i=1, i \neq n}^N \hat{\mathbf{g}}_n^H \hat{\mathbf{g}}_i x_i^d \\ &\quad + \alpha \sqrt{p_d} \sum_{i=1}^N \hat{\mathbf{g}}_n^H \mathbf{e}_i x_i^d + \alpha \hat{\mathbf{g}}_n^H \mathbf{n}^d + \hat{\mathbf{g}}_n^H \tilde{\mathbf{n}}^d, \end{aligned} \quad (15)$$

where  $\hat{\mathbf{g}}_i$  and  $\mathbf{e}_i$  are the  $i$ th columns of  $\hat{\mathbf{G}}$  and  $\mathbf{E}$ , respectively. The last four terms in (15) correspond to intra-cell interference, channel estimation error, receiver noise, and quantizer noise, respectively.

### C. Uplink Spectral Efficiency

Recall that the variance of elements of AWGN noise is 1, and the variance of elements of estimation error vector  $\mathbf{e}_i$  is  $\sigma_{\hat{g}_i}^2$ . According to the assumption of worst-case uncorrelated Gaussian noise [2] along with these variances, the uplink SE of the  $n$ th user can be given by

$$R_n = \frac{T - \tau}{T} \mathbb{E} \left\{ \log_2 \left( 1 + \frac{\alpha^2 p_d \|\hat{\mathbf{g}}_n\|^4}{\mathfrak{I}_{\hat{\mathbf{G}}}} \right) \right\}, \quad (16)$$

where  $\mathfrak{I}_{\hat{\mathbf{G}}} = \alpha^2 p_d \sum_{i \neq n} |\hat{\mathbf{g}}_i^H \hat{\mathbf{g}}_n|^2 + \alpha^2 p_d \|\hat{\mathbf{g}}_n\|^2 \sum_i \sigma_{\hat{g}_i}^2 + \alpha^2 \|\hat{\mathbf{g}}_n\|^2 + \alpha(1 - \alpha) \hat{\mathbf{g}}_n^H \text{diag}(p_d \mathbf{G} \mathbf{G}^H + \mathbf{I}) \hat{\mathbf{g}}_n$ .

Unfortunately, no efficient way is able to directly calculate the achievable SE in (16). Therefore, we derive an approximation that is presented in the following theorem.

**Theorem 1.** *For MRC receivers with MMSE channel estimates, the uplink SE (in bits/s/Hz) of the  $n$ th user in a quantized MIMO system, can be approximated as*

$$\mathcal{R}_n = \frac{T - \tau}{T} \log_2 \left( 1 + \frac{\tau \alpha^3 p_d p_t (M + 1) \beta_n^2}{\tau \Delta + (\alpha^2 + \tilde{\sigma}_t^2) \left( p_d \sum_{i=1}^N \beta_i + 1 \right)} \right) \quad (17)$$

where  $\Delta = \alpha^2 p_d p_t \beta_n \sum_{i=1}^N \beta_i + (1 - 2\alpha) \alpha^2 p_d p_t \beta_n^2 + \alpha^2 p_t \beta_n$ .

*Proof:* See Appendix.

Theorem 1 reveals the effect of several system parameters on the SE performance. In contrast to [5], Theorem 1 further involves the effect of channel estimate and also embraces [2] and [8] as special cases. Utilizing the expression in (17), we can investigate the optimal training length that should be allocated to maximize the uplink sum SE.

### III. OPTIMAL PILOT LENGTH

How much training is needed in MIMO systems has been studied extensively, e.g., [9], without considering the quantization effect. In contrast to the existing works, we consider this problem for the quantized MIMO setting. Specifically, we aim at analyzing the optimal pilot length  $\tau$  that maximizes the uplink SE:

$$\begin{aligned} \max_{\tau} \quad & \mathcal{S}(\tau) = \sum_{n=1}^N \mathcal{R}_n, \\ \text{s.t.} \quad & N \leq \tau \leq T, \end{aligned} \quad (18)$$

where  $\mathcal{S}(\tau)$  is the sum SE with  $\mathcal{R}_n$  given in (17), and the optimization variable  $\tau$  is an integer variable.

The optimal value  $\tau^*$  of (18) can be easily solved by a numerical algorithm since the set of feasible solutions is finite. To gain some insight into this optimization problem, we follow the methodology in [1] and assume that the powers of training and data transmission phases are fixed, i.e.,  $p_t = p_d$ . Additionally, we assume that all users exhibit the same large-scale fading, i.e.,  $\beta_n = \beta$ ,  $\forall n$ . Under these assumptions, (17) can be rewritten as

$$\mathcal{R}_0 = \frac{T - \tau}{T} \log_2 \left( 1 + \frac{a\tau}{b\tau + c} \right), \quad (19)$$

$$\begin{aligned} \text{where } a &= (M+1)\alpha^3 p_d^2 \beta^2, \\ b &= (N+1)\alpha^2 p_d^2 \beta^2 - 2\alpha^3 p_d^2 \beta^2 + \alpha^2 p_d \beta, \\ c &= (\alpha^2 + \tilde{\sigma}_t^2)(N p_d \beta + 1), \\ \tilde{\sigma}_t^2 &= \alpha(1 - \alpha)(1 + N p_d \beta). \end{aligned}$$

We thus obtain the sum SE

$$\mathcal{S}_0(\tau) = \left( 1 - \frac{\tau}{T} \right) N \log_2 \left( 1 + \frac{a\tau}{b\tau + c} \right). \quad (20)$$

To find the optimal solution of (20), we first provide the properties of  $\mathcal{S}_0(\tau)$  in the following lemma.

**Lemma 1.**  *$\mathcal{S}_0(\tau)$  is strictly concave in  $\tau$  for  $\tau > 0$ . In particular,  $\mathcal{S}_0(\tau)$  first increases and then decreases in the interval  $(0, \infty)$ .*

*Proof:* Through some basic calculations, we get that  $\mathcal{S}_0''(\tau)$ , the second derivatives of  $\mathcal{S}_0(\tau)$ , is less than 0. Moreover, we have  $\mathcal{S}_0'(0) > 0$ , and  $\mathcal{S}_0'(\tau) \rightarrow -N \log_2(1 + \frac{a}{b})/T < 0$  as  $\tau \rightarrow \infty$ . In addition,  $\mathcal{S}_0'(\tau)$  is monotonically decreasing with  $\tau$ . From these properties, we can deduce that  $\mathcal{S}_0$  first increases and then decreases. ■

The lemma above treats  $\tau$  as a continuous variable in  $(0, \infty)$ . However, the pilot length must be a positive integer in the interval  $[N, T]$ . By introducing this constraint, we have the following proposition.

**Proposition 1.** *The optimal pilot length  $\tau^*$  of the problem (18) obeys the following properties:*

- if  $\mathcal{S}_1'(N) > 0$ , then  $\tau^*$  can be obtained by Newton's method and then rounded to the nearest integer;
- if  $\mathcal{S}_1'(N) \leq 0$ , then  $\tau^* = N$ .

*Proof:* From Lemma 1, we know that  $\mathcal{S}_0(\tau)$  is concave in  $\tau$  for  $\tau > 0$ . Since  $\mathcal{S}_0'(T) = -\frac{N}{T} \log_2 \left( \frac{aT}{bT+c} + 1 \right) < 0$ , we obtain that if  $\mathcal{S}_0'(N) > 0$ , then  $\mathcal{S}_0(\tau)$  first increases and then decreases in the interval  $[N, T]$ . Under this condition,  $\mathcal{S}_0(\tau)$  reaches its maximum value at  $\hat{\tau}^*$  which satisfies  $\mathcal{S}_0'(\hat{\tau}^*) = 0$ . The value of  $\hat{\tau}^*$  can be obtained by Newton's method, and then  $\tau^*$  is rounded to its nearest integer. On the other hand, if  $\mathcal{S}_0'(N) \leq 0$ ,  $\mathcal{S}_0(\tau)$  is a monotonically decreasing function over the range  $N \leq \tau \leq T$ . Therefore, the optimal  $\tau^*$  is  $N$ . ■

Proposition 1 shows that the optimal pilot length is dependent on the value of  $\mathcal{S}_0'(\tau)$ , which relates to many system parameters such as the transmit power and number of quantization bit.

### IV. NUMERICAL RESULTS

In the simulations, we consider a hexagonal cell with radius of 1000 meters. The users are distributed randomly and uniformly over the cell, with the exclusion of a central disk of radius  $r_h = 100$  meters. The large-scale fading is modeled as  $\beta_n = z_n / (r_n / r_h)^v$ , where  $z_n$  is a log-normal variable with standard deviation  $\sigma_{\text{shadow}}$ ,  $r_n$  is the distance between the  $n$ th user and BS, and  $v$  is the path loss exponent. Note that  $\beta_n$  is fixed once the  $n$ th user is dropped in the cell and the expectation is taken over the small-scale fading coefficients. In all the following experiments, we assume that  $\sigma_{\text{shadow}} = 8$  dB,

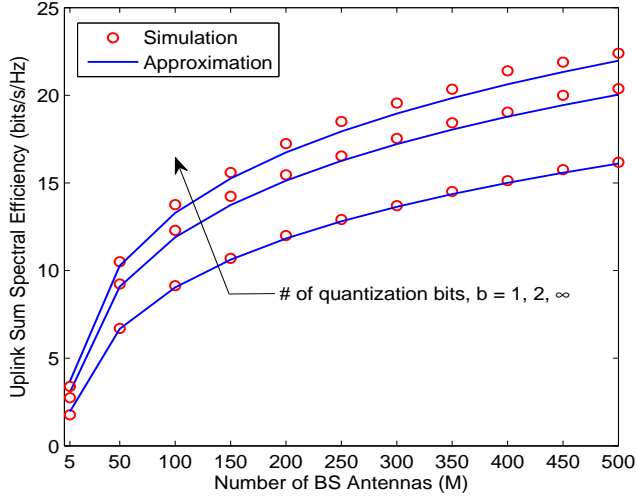


Fig. 1. Uplink sum SE per cell versus the number of BS antennas  $M$ , where  $N = 10$ ,  $\tau = 10$  and  $p_d = 10$  dB.

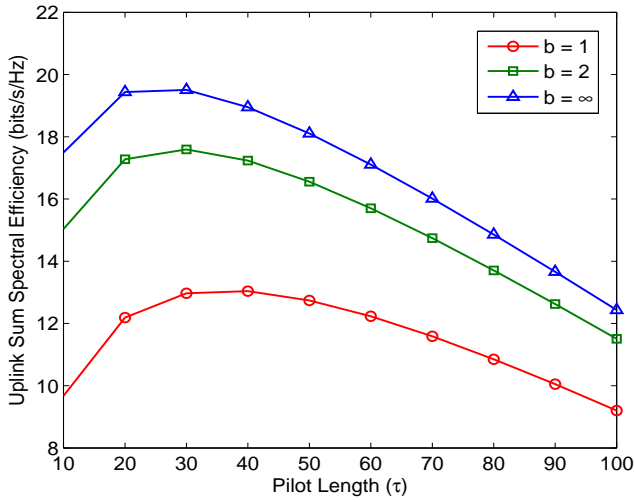


Fig. 2. Uplink sum SE per cell versus the pilot length  $\tau$ , where  $M = 100$ ,  $N = 10$ , and  $p_d = -10$  dB.

$v = 3.8$ ,  $p_d = p_t$ , and the coherence time  $T = 196$  in symbols [8]. Since the noise variance is set as unit in (1), the SNR of the system is defined as  $\text{SNR} = p_d$  and given in dB.

We first validate the accuracy of our proposed approximation in Theorem 1. In Fig. 1, the simulated sum SE in (16) is compared with its corresponding analytical approximation in (17). Results are presented for three different quantizers with 1, 2, and  $\infty$  bits, respectively, for  $N = 10$  users and  $p_d = 10$  dB. In all cases, a precise agreement between the simulation results and our analytical results can be observed.

Next, we analyze the sum SE with respect to the pilot length  $\tau$ . Given the tightness between the simulated values and the approximations, we adopt the analytical approximation (17) in the following discussion. In this experiment, we assume all users have the same large scale fading,  $\beta = 1$ . Fig. 2 shows the corresponding sum SE for three different values of quantization bits. These curves increase first and then decrease,

TABLE I  
THE OPTIMAL PILOT LENGTH  $\tau^*$  FOR DIFFERENT TRANSMIT POWER  $p_d$  AND QUANTIZATION BITS  $b$ .

$b \backslash p_d$ (dB)	-10	0	10	20	30	40
1	36	22	19	19	19	19
2	28	13	10	10	10	10
3	26	10	10	10	10	10
$\infty$	25	10	10	10	10	10

which agrees with our argument in Lemma 1. This property can be understood by the trade-off between the channel estimation accuracy and the data transmission time. Although all curves show the same trend, the optimal pilot length  $\tau^*$  and the slope of the curves are different. In particular, we find that  $\tau^*$  decreases with increasing  $b$ . This property implies that a longer training length is required for low-resolution quantizers.

The optimal pilot length for different  $b$  and  $p_d$  are given in Table I. We see that  $\tau^*$  varies with  $b$  and  $p_d$ , and this is consistent with Proposition 1. Interestingly, we notice that  $\tau^*$  for  $b = 1$  is quite different from the others since  $\tau^*$  converges to 19 instead of 10 (the number of users). This result implies that for the extreme case with one-bit quantizers, the increase of transmit power cannot unlimitedly improve the accuracy of the channel estimate. In this case, the pilot length must be greater than the number of users even in the high SNR regime.

## V. CONCLUSION

We investigated the optimal pilot length for quantized massive MIMO systems with coarse resolution quantizers. A tight approximate expression for the achievable uplink SE was derived. Using the analytical expression, we then deduced the optimal pilot length which maximizes the sum SE. Results shows that a longer training length is required for coarse resolution quantizers. In particular, for MIMO receivers with one-bit quantizers, increasing the transmit power cannot sufficiently improve the channel estimation such that longer training interval length is required.

## APPENDIX

Applying [8, Lemma 1], we begin by approximating  $R_n$  with

$$R_n \approx \mathcal{R}_n = \frac{T - \tau}{T} \log_2 \left( 1 + \frac{p_d \alpha^2 \mathbb{E} \{ \|\mathbf{g}_n\|^4 \}}{\mathbb{E} \{ \mathcal{I}_{\hat{\mathbf{G}}} \}} \right) \quad (21)$$

where

$$\begin{aligned} \mathbb{E} \{ \mathcal{I}_{\hat{\mathbf{G}}} \} &= \alpha^2 p_d \sum_{i=1, i \neq n}^N \mathbb{E} \{ |\hat{\mathbf{g}}_n^H \hat{\mathbf{g}}_i|^2 \} + \alpha^2 p_d \sum_{i=1}^N \mathbb{E} \{ \|\hat{\mathbf{g}}_n\|^2 \} \sigma_{\hat{\mathbf{g}}_i}^2 \\ &+ \alpha^2 \mathbb{E} \{ \|\hat{\mathbf{g}}_n\|^2 \} + \alpha (1 - \alpha) \mathbb{E} \{ \hat{\mathbf{g}}_n^H \text{diag}(p_d \mathbf{G} \mathbf{G}^H + \mathbf{I}_M) \hat{\mathbf{g}}_n \}. \end{aligned} \quad (22)$$

Although the true distribution of  $\hat{\mathbf{g}}_n$  cannot be determined due to the quantization, we approximate it as Gaussian as in [6], and rewrite  $\hat{\mathbf{g}}_n$  as

$$\hat{\mathbf{g}}_n = \sigma_n \mathbf{h}_n, \quad (23)$$

where  $\mathbf{h}_n \sim \mathcal{CN}(0, \mathbf{I}_N)$ , and  $\sigma_n^2$  is the variance of  $\hat{\mathbf{g}}_n$ . According to the orthogonality principle, we obtain

$$\sigma_n^2 = \beta_n - \sigma_{\hat{\mathbf{g}}_n}^2 = \frac{\tau \alpha^2 p_t \beta_n^2}{\tau \alpha^2 p_t \beta_n + \alpha^2 + \tilde{\sigma}_t^2}, \quad (24)$$

where  $\sigma_{\hat{\mathbf{g}}_n}^2$  is given in (11). Then, following the proof in [5], we can obtain the desired result in (17).

#### REFERENCES

- [1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590-3600, Nov. 2010.
- [2] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436-1449, Apr. 2013.
- [3] J. Mo and R. W. Heath Jr., "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5498-5512, Oct. 2015.
- [4] O. Orhan, E. Erkip, and S. Rangan, "Low power analog-to-digital conversion in millimeter wave systems: Impact of resolution and bandwidth on performance," preprint, available in <http://arxiv.org/abs/1502.01980>.
- [5] L. Fan, S. Jin, C.-K. Wen, and H. Zhang, "Uplink achievable rate for massive MIMO systems with low-resolution ADC," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2186-2189, Dec. 2015.
- [6] Q. Bai, U. Mittmann, A. Mezghani, and J. A. Nossek, "Minimizing the energy per bit for pilot-assisted data transmission over quantized channels," in *Proc. IEEE PIMRC*, Sept. 2014.
- [7] Q. Bai, A. Mezghani, J. A. Nossek, "On the optimization of ADC resolution in multi-antenna systems," in *Proc. IEEE ISWCS*, Aug. 2013.
- [8] Q. Zhang, S. Jin, K. K. Wong, H. B. Zhu, and M. Matthaiou, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 966-981, Oct. 2014.
- [9] H. Q. Ngo, M. Matthaiou, and E. G. Larsson, "Massive MIMO with optimal power and training duration allocation," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 605-608, Dec. 2014.